



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Dynamics of uncertainty in geological interpretation

Citation for published version:

Polson, D & Curtis, A 2010, 'Dynamics of uncertainty in geological interpretation', *Journal of the Geological Society*, vol. 167, no. 1, pp. 5-10. <https://doi.org/10.1144/0016-76492009-055>

Digital Object Identifier (DOI):

[10.1144/0016-76492009-055](https://doi.org/10.1144/0016-76492009-055)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Journal of the Geological Society

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



This is the Author's final draft or 'post-print' version. The final published version has been accepted by the Journal of the Geological Society.

Cite As:

Polson, D., & Curtis, A. (2010). Dynamics of uncertainty in geological interpretation. *Journal of the Geological Society*, 167(1), 510. 10.1144/0016-76492009-055

DOI: 10.1144/0016-76492009-055

Dynamics of uncertainty in geological interpretation

Author:

Debbie Polson and Andrew Curtis
Osborn.

Address for correspondence:

Debbie Polson
School of Geosciences,
University of Edinburgh,
The King's Buildings
West Mains Road,
Edinburgh

,
UK

EH9 3JW

Email: dpolson@staffmail.ed.ac.uk

Dynamics of uncertainty in geological interpretation

Debbie Polson* and Andrew Curtis

University of Edinburgh, School of GeoSciences, Edinburgh EH9 2JT
Also at ECOSSE (Edinburgh Collaborative of Subsurface Science and Engineering)
corresponding author (e-mail: dpolson@staffmail.ed.ac.uk)

No. of words: 2403; No. of figures: 3

Running title: uncertainty in geological interpretation

Abstract:

Interpretation of geological data is based on both personal judgement and previous experience of related scenarios. In combining such information geologists employ heuristics (rules of thumb), and are therefore subject to biases that are well known in cognitive psychology and are common to all expert judgements. Here we analyse dynamic uncertainty in an evolving geological interpretation. Through a well-designed elicitation process we show how the inclusion of multiple experts influences interpretational bias. In particular, group convergence of opinion is observed, and we show how this can be differentiated from ‘herding’ behaviour similar to that observed in economic bubbles by forcing a consensus to be reached. Thus we can identify when and why the judgement of a single geological expert should be treated with caution. This process can be applied to any geological interpretational scenario.

Geologists are often required to make judgements and interpretations in situations of uncertainty where data are inadequate to fully constrain any particular interpretation. Any interpretation by an expert is then dependent on the prior knowledge and experience of that expert, and hence the result is both subjective and qualitative in nature. The geological prior information employed is difficult to assess or quantify as experience and knowledge vary from expert to expert, as do the methods an expert may use to generalise and categorise information (Rankey & Mitchell 2003; Wood & Curtis 2004; Curtis & Wood 2004a; Bond et al., 2007)

This prior information may be thought of as a prior probability that an expert places on each hypothesis. As new information, (e.g. data) becomes available this initial probability is updated by combining it with the data to produce a new, ‘posterior’ probability of each hypothesis. However, this approach assumes that the way in which an expert forms and updates their beliefs follows some rational, ordered approach. Research has shown that all experts are subject to biases when making probabilistic assessments which result in inaccurate and uncertain judgements (e.g. Baddeley et al, 2004, O’Hagan et al, 2006). For example, numerous studies have shown that all individuals find it difficult to assign numerical probabilities to judgements (e.g. Kahneman et al., 1982, Anderson,

1998). This is because heuristics (rules of thumb) are used to assess probability, and these introduce bias. Group interaction has the advantage of allowing knowledge and experience to be shared amongst the experts. It provides a method of aggregating individual opinions (e.g. Phillips, 1999) and evidence suggests that group interaction can reduce some effects of individual bias (Snizek, 1992). However this can also introduce other group biases (Snizek, 1992).

This paper investigates uncertainty in geological interpretation by individual and multiple experts. It describes an elicitation process designed to demonstrate individual bias, and the affect of group dynamics both on the final interpretations, and on the perceptions of individual experts. Expert elicitation theory and practice have been investigated (e.g. see Bonano et al, 1991) and used in numerous studies, including in the earth and environmental sciences (Morgan et al, 2001; Curtis & Wood 2004a; Arnell et al, 2005; Lowe et al, 2006; Bond et al., 2007 and Ye et al, 2008). However a well structure and well managed elicitation process is essential to avoid group biases such as overconfidence, to prevent the group becoming dominated by opinion over knowledge and to ensure that the expertise of all individuals is recognised with no single individual dominating the group by force of personality (O’Hagan, 2006).

The results show that individual judgements can be contradictory, and that group interaction radically alters individual perceptions. They demonstrate that a group consensus may not reflect the opinions of all constituent and consenting experts, and overall the results show that any probabilistic assessment, whether from individuals or a group, should be elicited with a carefully structured process such as the one we propose here if the results are to be properly valued and understood.

Context

A set of 2D seismic lines and contextual geological data of the Firth of Forth has been interpreted by experts on at least three occasions to produce 3-D subsurface geological models. The Firth of Forth is in the Midland Valley of Scotland (MVS). The Midland Valley is an 80 km wide and >150 km long NE-trending graben structure consisting of a series of sedimentary basins formed between 390-280 million years ago in Devonian, Carboniferous and Permian times. These form a vertical succession of 5 km of mainly fluvio-lacustrine and marginal marine sedimentary rocks (Ritchie et al., 2003). The rock types are a variety of intercalated mudstones, siltstones, sandstones, coals, limestones, and extrusive and intrusive igneous rocks. There are a series of NNE trending anticlinal and synclinal folds with numerous large faults.

Geological Data

The seismic dataset consisted of thirty-three, 2-D seismic reflection profiles in the Firth of Forth in Scotland (see Fig 1 for an example). The lines are arranged in an approximate 1 km x 1 km grid and have a combined length of approximately 600 km. A number of commercial wells and boreholes were used to calibrate the data. These include the 25/26-1 well drilled by Conoco in 1990 which reached a depth of 2009.5m and shallow stratigraphic boreholes drilled by the Institute of Geological Studies (now the British Geological Survey) and the National Coal Board (Thomson, 1978). Structural contour information from mine-working operations, published refraction work and evidence from field based studies also contributed to the dataset. For a full description see Ritchie et al. (2003) and Underhill et al. (2007).

Interpretation of the Geological Data

From the geological models of the region based on this data, three features were identified for the elicitation exercise. These included a large fault, a sandstone formation and an overlaying mudstone formation.

One interpretation of the data includes a fault running north-south referred to as the Mid-Forth Fault (Ritchie et al., 2003) which is absent from later interpretations (e.g. Underhill et al., 2007).

A particular formation, the Knox Pulpit sandstone formation is known to exist to the North of the Firth of Forth from borehole and outcrop data. Palaeogeographic reconstructions suggest that this formation is likely to be present at depth within the Firth of Forth (Browne et al., 1987), but this remains unproven.

Above the Knox Pulpit Formation is the mainly siltstone and mudstone Ballagan Formation which is known to exist both North and South of the Firth of Forth (Mitchell and Mykura, 1962) but is also unproven in the Firth of Forth.

Method

An elicitation exercise was carried out with the four experts to investigate the current level of uncertainty in individual expert interpretations of this dataset, and the effects of group interaction.

The experts were asked to assess the probability that three structures in the earth model existed in a subregion beneath the Firth of Forth: the Knox Pulpit sandstone formation (a potential reservoir), the Ballagan Formation above the reservoir, and the Mid-Forth Fault, hereafter referred to as reservoir, seal and fault respectively. They also assigned lower- and upper-bound probabilities to their best-guess, allowing each expert to make an assessment of their own uncertainty.

A six step elicitation exercise took place: (1) Information was elicited from each expert independently before the group session. (2) At the start of the group session they were alerted to the various common biases in expert judgements. (3) The experts repeated the assessment individually. (4) These individual assessments were shared amongst the group with each expert asked to explain their reasoning. (5) This led to group discussion through which the experts were prompted to reach a consensus assessment. (6) Finally the experts again repeated assessments individually.

Results

Figs 2, 3 and 4 show the probability distribution (p.d.'s) (from linear interpolation of lower, best-guess and upper bounds) for each structure. These are normalised distributions which give the approximate probability of percentage, p , lying within a particular range by integrating the probability density function, $f(p)$, over the given range. Expert E4 was not available prior to the session and so has no pre-session results. During the introduction stage of the group session it became clear that expert E1 had assessed the wrong fault. This resulted in a radically different distribution at session start compared to that pre-session (compare Figs 4(a) and 4(b)). During discussions regarding the existence of a seal, it was revealed that this expert had also incorrectly assumed that instead of assessing the probability that the seal exists they should account for the quality of the seal. Figs 2(b) and 2(c) show the change in perception of expert E1 from the start to the end of the session with the assessment at the end showing a significant change.

Fig. 2(a) shows pre-session distributions for the probability of the seal. Expert E1 misunderstood what was being assessed and experts E2 and E3 reached different conclusions from each other about the probability of the seal's existence, with their distributions overlapping only at their upper and lower bounds respectively.

Fig. 3(a) shows pre-session results for the existence of the reservoir. While the distributions of experts E2 and E3 mostly overlap, expert E1 shows significant disagreement. Expert E1 is more certain of the reservoirs existence (higher probability) and is more confident of this assessment (narrow p.d.).

Fig. 4(a) shows pre-session results for the existence of the fault. Expert E1 has misunderstood which fault they are assessing while opinions of experts E2 and E3 overlap. However, at the start of the session shown in Fig. 4(b), expert E3 has radically changed their assessment resulting in distributions from experts E2 and E3 that are mutually exclusive. This shows that individual experts can make interpretations based on the same set of data which are contradictory with others and with themselves, even taking into account the expert's estimation of their own uncertainty.

Individual assessments at the end of the elicitation session (Figs 2(c), 3(c) and 4(c)) show a clear shift towards the consensus distribution compared to pre-

session and start of session assessments. The experts have individually converged towards a group position during the elicitation process. Expert E1 in particular shows a dramatic change with respect to the reservoir for which there was no misunderstanding of the task: the distribution in Fig. 3(a) shows a high degree of confidence in the existence of the reservoir with low uncertainty before the session. By the end of the session, this expert has radically changed their opinion with a final assessment, shown in Fig. 3(c) that excludes their initial assessment entirely.

Results from parts (a), (b) and (c) individually were combined by averaging the lower, best-guess and upper bounds. Where there was a misunderstanding, results were excluded. The average individual assessments as shown in Figs 2(d), 3(d) and 4(d) for the seal, reservoir and fault respectively, show a clear convergence towards a consensus during the elicitation process. While the average distribution from the end of the session differs from the consensus, it shows a significant shift towards the consensus for all three structures.

Combining the distributions for the seal, reservoir and fault gives the joint probability distribution for the existence of all three features at once (if the fault was impermeable), this might represent the p.d. of the existence of a viable bounded reservoir. Fig 5 shows the distributions for the seal, reservoir and fault combined into a single probability distribution for the individual pre-group elicitation as well as the average of the pre-group elicitation, start of elicitation, and end of elicitation and the combined consensus distributions reached by the experts during the session. Fig 5(a) shows the simplified method of averaging the experts' distributions by averaging the lower-bounds, best-guess and upper-bounds elicited from the individual experts. Fig 5(b) shows the combined distributions produced by averaging the individual expert distributions for the seal, reservoir and fault. This method is more robust as it shows the full width of all four experts' individual distributions. Fig 5(c) shows the combined individual expert pre-session distributions.

Figs 5 (a) and (b) both show the joint probability increasing as the elicitation progressed with the peak of the end of session distribution greater than the pre-session and start of session distributions. The distribution from the end of the session is also wider than the earlier distributions as the experts become less confident through the elicitation process. In both cases the end of session distribution is similar to the consensus. From the peak of the consensus

distribution, the most likely probability that all three structures exist is ~25%, with an upper limit of ~60%. Figure 5(c) shows the variation between experts before the elicitation session with experts E1 and E2 producing distributions that are mutually exclusive.

Discussion

The results of the elicitation demonstrate clearly for the first time the dynamics of uncertainty in the interpretation of geological data by teams of multiple experts. The distributions of individual experts at pre-session and at start of session highlight the potential for experts to reach contradictory conclusions from the same data, even given their own estimation for their potential for error, for example experts E1 and E3 in Figs. 2(a) and 3(a). This reflects the subjective nature of the prior knowledge and experience of experts, which influences their interpretation.

In real life lone experts will regularly be required to make judgements in situations of uncertainty while having more than three experts involved in such an assessment is rare. However the results of this study highlight the risk associated with using single experts due to the potential for disagreement as highlighted in Fig 3(a). Had only expert E1 been asked to assess the probability of the existence of the reservoir, the result would be a confident assessment of a high probability of existence. However the other two experts elicited show far less certainty about the reservoirs existence and are far less confident in their answers, as is expert E1 by the end of the process (Fig 3(c)).

Of particular concern is the potential for misunderstandings and incorrect assumption which may remain undetected. Misunderstandings affected two of the three pre-session assessments for one expert despite every effort having been made to explain the task. Other studies (e.g. Phillips, 1999) have also shown that when required to make an assessment, experts may make different prior assumptions resulting in radically different assessments. It is only through the structured elicitation process that these assumptions and misunderstanding were detected. It is possible that in real life interpretations, assumptions and misunderstandings commonly occur in expert judgements but remain undetected.

Group elicitation allows knowledge and expertise to be shared amongst experts and can reduce bias. The results from this study clearly demonstrate pooling of information, resulting in changes in individual opinions. This process is observed for the fault, reservoir and seal suggesting a systematic process that results in a convergence of individual positions towards a group consensus. The first two stages of the structured elicitation process measure the existing beliefs of the participants. Through latter stages the elicitation process itself acts to create new ideas and insights, meaning that the process itself actually creates beliefs. As the session progressed, individual experts became less certain of their assessment resulting in wider distributions. This highlights the potential for the elicitation process to reduce well-known cognitive biases, in this case expert overconfidence.

For the seal and reservoir where there was agreement amongst the group and consensus was relatively easy to reach; the final average individual assessments, Figs 2(d) and 3(d) respectively, closely matched the consensus. The consensus distribution also bounded or nearly bounded the range of individual opinions, as shown in Figs 2(c) and 3(c) for the seal and reservoir respectively, encapsulating the combined knowledge of the group.

For the fault, agreement was not as strong and the consensus was partially forced (as is typical in a decision-making scenario). The final consensus assessment does not reflect the combined knowledge and opinions of all individuals in the group. In this case 'herding', where the group has been lead by or followed one individual (Keynes, 1921, 1937), is clearly observed and for the first time quantified: the group has moved towards a single member's opinion, in this case expert E1, with the consensus distribution primarily reflecting the views of this one individual as shown in Fig. 4(c). In order to reach a consensus, individual members of the group have agreed to a distribution in which they did not truly believe. One individual in particular, expert E4, gave a final individual distribution that was significantly outside of the group consensus to which they had previously agreed.

Where one expert disagrees with the majority of the group, care must be taken to ensure that the potentially important views and information held by this individual are not neglected. It is therefore of significant concern when the final consensus does not span such opinions and in such cases where a consensus is not easily forthcoming, mathematical aggregation of individual assessments may better

represent the full scope of the group knowledge and expertise. Nevertheless, the combination of a consensus p.d. with post-consensus, individual p.d.s is a demonstrably a key tool to identify such herding behaviour.

Due to issues with herding and the consensus distribution for the fault, it is likely that the results for the seal and reservoir are relatively more reliable. However without the structured elicitation process which allowed both misunderstandings and herding to be identified, and herding to be differentiated from mere convergence of views, we would not have been able to reach this conclusion. It is only by tracking the evolution of individual beliefs and comparing such dynamics with final consensus distributions that we were able to identify herding behaviour and were alerted to problems with the very nature or meaning of a 'consensus' distribution. This method should therefore be applied in a wide variety of interpretation tasks in future.

Using structured, expert elicitation it is possible to produce a probability distribution that a reservoir has the basic required characteristics for security based on the initial assessment of data by experts. In this case we assess the probability that the reservoir exists and that it has a seal and fault. If we assume that this is a sealing fault required for reservoir security, the joint probability gives the probability that the site has the minimum required features. This however would be a maximum probability based on the existing data as other features and characteristics would no doubt also be required. This assessment can then be used to inform decisions regarding further exploratory work and data acquisition activities. The distributions produced by individual experts before the elicitation session show significant variability compared to the consensus and averaged individual end of session distributions highlighting the risk of using individual experts and the value of using a structured elicitation process.

References

- Anderson, J. L. 1998. Embracing uncertainty: the interface of Bayesian statistics and cognitive psychology. *Conservation Ecology*, 2, 2. [Online] World Wide Web Address: <http://www.consecol.org/vol2/iss1/art2>.
- Arnell, N.W., Tompkins, E.L., Adger, A.N., 2005. Eliciting information from experts on the likelihood of rapid climate change. *Risk Analysis* **25** (6), 1419–1431.
- Baddeley, M. A., Curtis, A. and Wood, R. A. 2004. An introduction to prior information derived from probabilistic judgements: elicitation of knowledge, cognitive bias and herding. In: Curtis, A. and Wood, R. A. (2004b), **239**, p. 15-27.
- Bond, C. E., Gibbs, A., Shipton, Z. K., and Jones, S., 2007. What do you think this is? "Conceptual uncertainty" in geoscience interpretation. *GSA Today*, **17**, no. 11, p. 4-10.
- Bonano, E.J., Apostolakis, G.E., 1991. Theoretical foundation and practical issues for using expert judgments in uncertainty analysis of high-level radioactive waste disposal. *Radioactive Waste Management and the Nuclear Fuel Cycle*, **16** (2), 137-159.
- Browne, M.A.E, Robins, N S, Evans, R B, Monro, S.K. and Robson, P.G. 1987. The Upper Devonian and Carboniferous sandstones of the Midland Valley of Scotland. Investigation of the Geothermal Potential of the UK, British Geological Survey.
- Curtis, A. and Wood, R. 2004a. Optimal elicitation of probabilistic information from experts. In: Curtis, A. and Wood, R. A. (2004b), **239**, p. 127-145.
- Curtis, A. and Wood, R. 2004b. Geological Prior Information: Informing Science and Engineering. London, Geological Society, London Special Publications.
- Kahneman, D., Slovic, P., and Tversky, A. (eds), 1982. *Judgement under Uncertainty: Heuristics and Biases*. Cambridge, Cambridge University Press.
- Keynes, J. 1921. *A Treatise on Probability*. Macmillan, London.
- Keynes, J. 1937. The general theory of employment. *Quarterly Journal of Economics*, **51**, p. 209-223.
- Lowe, T. D. and Lorenzoni, I. 2007. Danger is all around: Eliciting expert perceptions for managing climate change through a mental models approach. *Global Environmental Change*, **17**, 131-146.
- Ming Ye, Karl F. Pohlmann, Jenny B. Chapman, 2008. Expert elicitation of recharge model probabilities for the Death Valley regional flow system. *Journal of Hydrology* **354**, 102- 115.
- Mitchell, G H, and Mykura, W. 1962. The geology of the neighbourhood of Edinburgh. Third edition. Memoir of the Geological Survey, Sheet 32 (Scotland).
- Morgan, M. G., Pitelka, L. F. and Shevliakova, E. 2001. Elicitation of expert judgments of climate change impacts on forest ecosystem. *Climatic Change* **49**: 279–307, 2001.
- O'Hagan, A., Buck, C. E, Daneshkhan, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rakow T. 2006. *Uncertain Judgements: Eliciting Expert's Probabilities*. Wiley.
- Phillips, L. D., 1999. Group elicitation of probability distributions: are many heads better than one? In: Shanteau J., Mellors, B. and Schum, D. (eds) *Decision Science and Technology: Reflections on the contributions of Ward Edwards*, Norwell MA: Kluwer Academic Publishers, p. 313-330.
- Rankey, E.C., and Mitchell, J.C., 2003. That's why it's called interpretation: The role of horizon uncertainty on seismic attribute analysis: *The Leading Edge*, p. 820-828.
- Ritchie, J D., Johnson, H., Browne, M.A.E., Monaghan, A.A. 2003. Late Devonian-Carboniferous tectonic evolution within the Firth of Forth, Midland Valley; as revealed from 2D seismic reflection data. *Scottish Journal of Geology* **39**, 121-134.
- Snizek, J. A. 1992. Groups under uncertainty: An examination of confidence in group decision making. *Organizational Behaviour and Human Decision Processes*, **52**, p. 124-155.
- Thomson M E, 1978. IGS studies of the geology of the Firth of Forth and its approaches. Report of the Institute of Geological Sciences, No.77/17.
- Underhill, J.R., Monaghan, A.A. and Browne, M.A.E. 2008. Controls on Structural Styles, Basin Development and Petroleum

Prospectivity in the Midland Valley of Scotland. *Journal of Marine and Petroleum Geology*, **25**, pp. 1000-1022.

Wood, R. and Curtis, A. 2004. Geological prior information and its application to geoscientific problems. In: Curtis, A. and Wood, R. A. (2004b), **239**, p. 1-14.

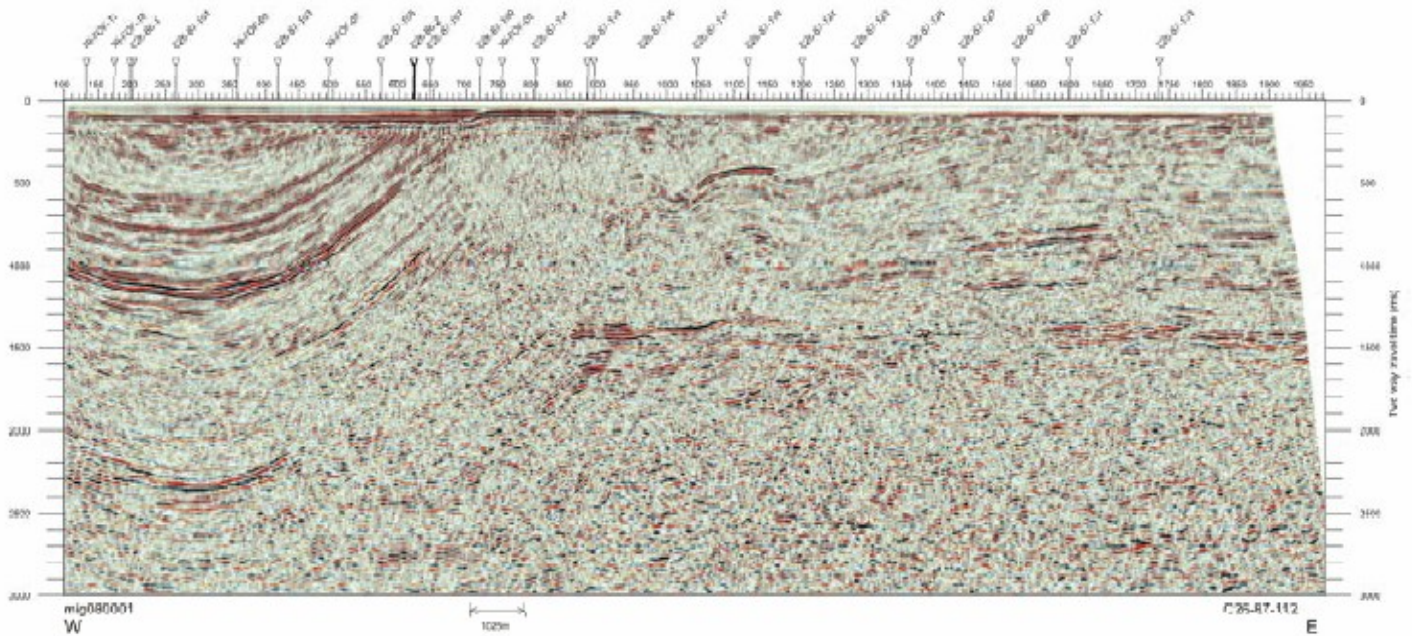


Fig. 1. Example of Firth of Forth seismic data. Line C26_87_112 oriented SW-NE through the 25/26-1 tied well. BGS©NERC (IPR/114-29DR), seismic data shown with permission of Phoenix Data Solutions.

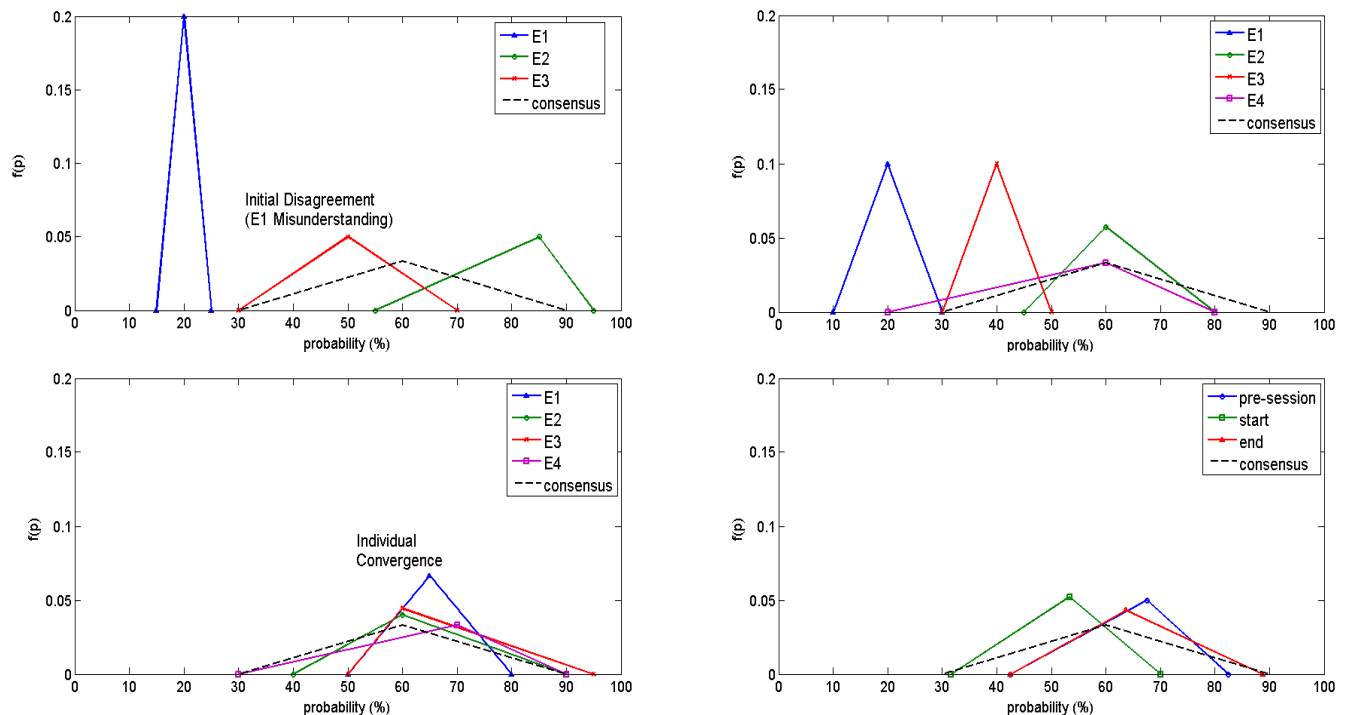


Fig. 2. Probability density function for the existence of the seal from individual (solid) and consensus (dashed) assessment of the lower, upper and best-guess probabilities labelled with the interpretation of the results. The probability of the percentage, p , falling within a given range is given by the integral of $f(p)$ over the range. (a) pre-session assessments showing initial disagreement amongst experts and a misunderstanding for expert E1 (b) start of session assessments, (c) end of session assessments showing the individual experts converging towards to group consensus, and (d) average. E1 to E4 represent the four experts.

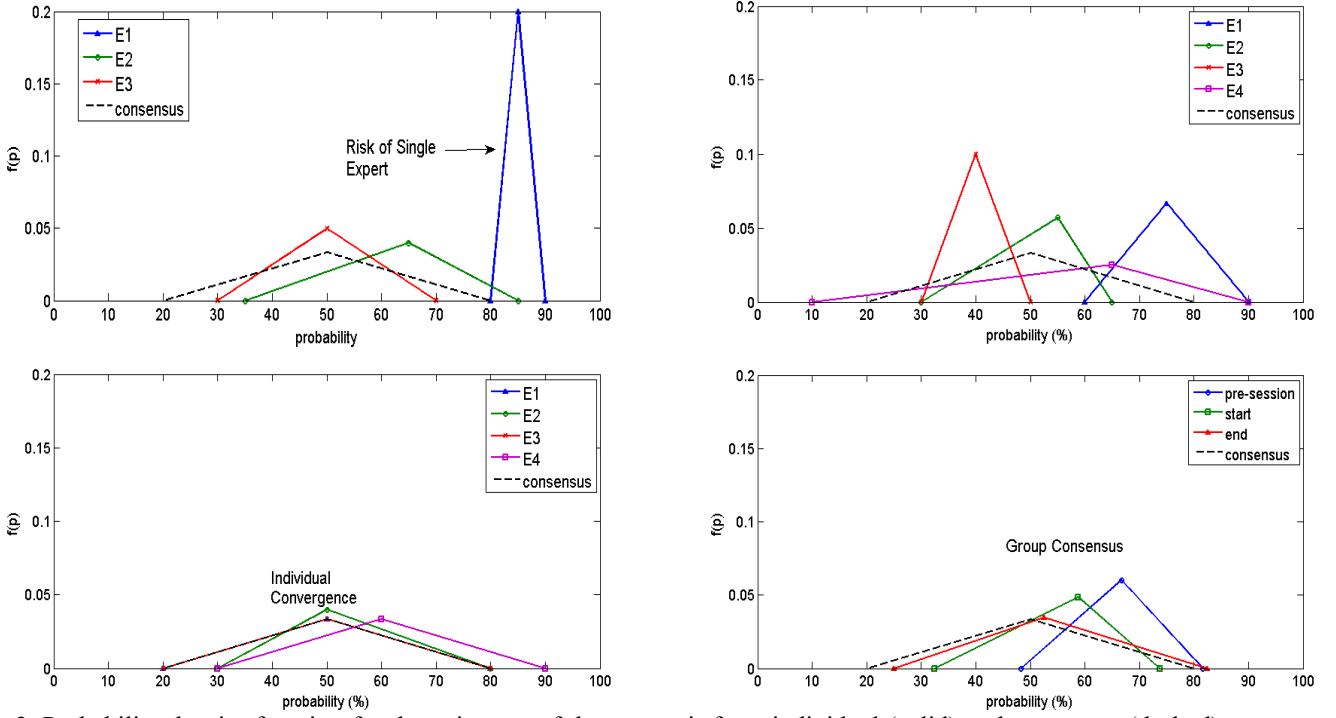


Fig. 3. Probability density function for the existence of the reservoir from individual (solid) and consensus (dashed) assessment of the lower, upper and best-guess probabilities labelled with the interpretation of the results. The probability of the percentage, p , falling within a given range is given by the integral of $f(p)$ over the range. (a) pre-session assessments showing the potential risk of eliciting the judgement of only one expert which may differ radically from the judgements made by other experts, (b) start of session assessments, (c) end of session assessments showing the individual experts converging towards the group consensus and (d) average. E1 to E4 represent the four experts showing the group convergence through the elicitation process.

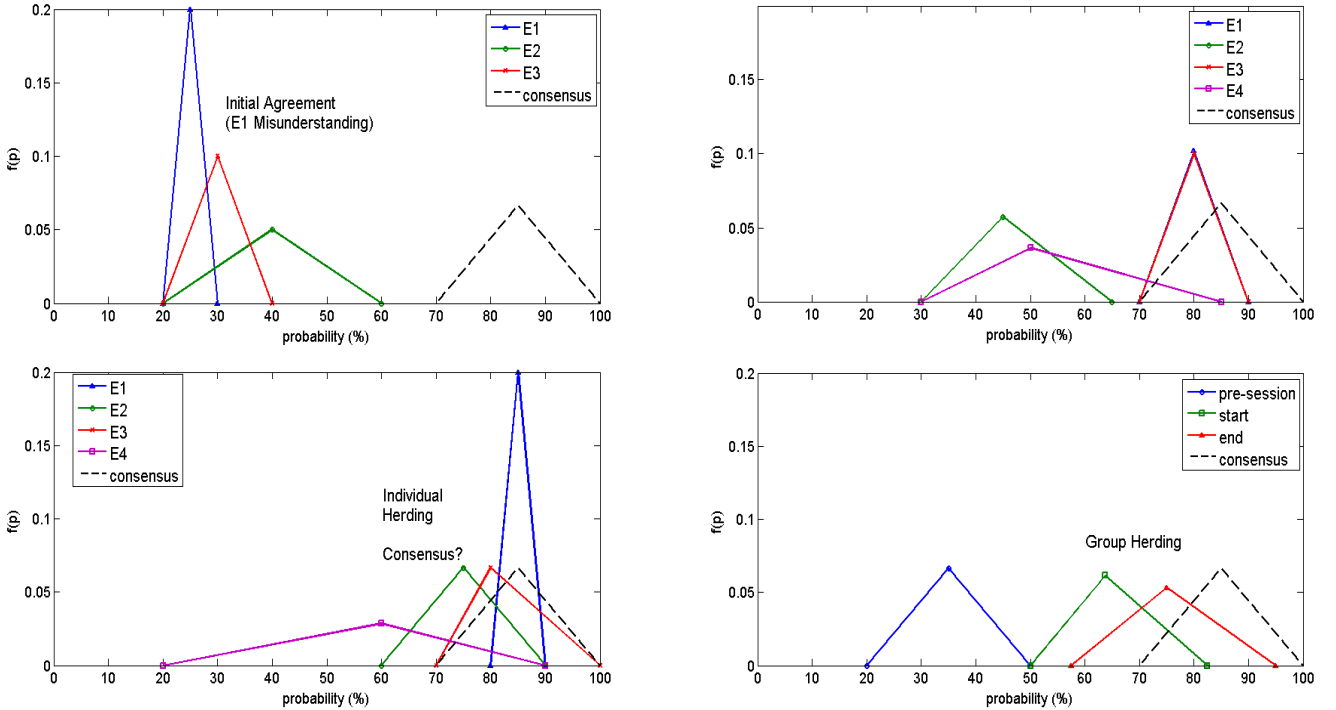


Fig. 4. Probability density function for the existence of the fault from the individual (solid) and consensus (dashed) assessment of the lower, upper, and best-guess probabilities labelled with the interpretation of the results. The probability of the percentage, p , falling within a given range is given by the integral of $f(p)$ over the range. (a) pre-session assessments showing initial agreement between experts one misunderstanding from expert E1, (b) start of session assessments, (c) end of session assessments showing the individual experts herding around expert E1, and (d) average. E1 to E4 represent the four experts showing group herding through the elicitation process

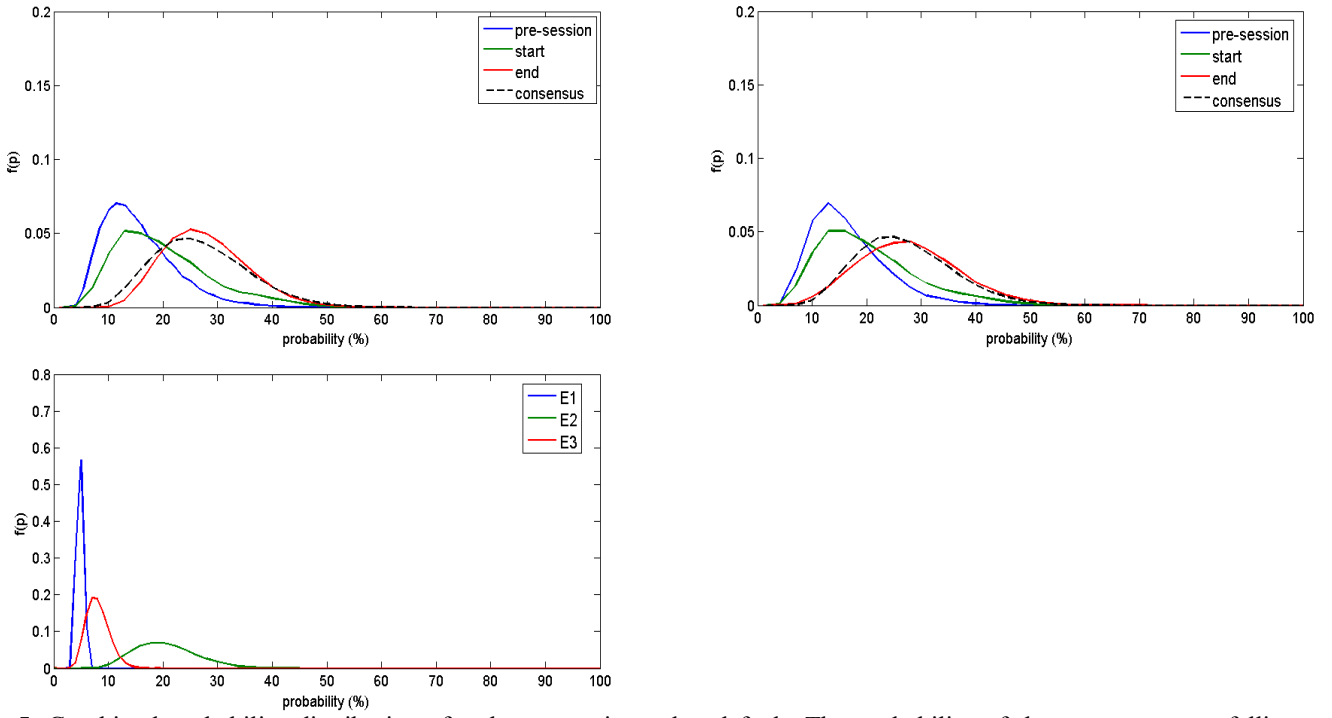


Fig. 5. Combined probability distributions for the reservoir, seal and fault. The probability of the percentage, p , falling within a given range is given by the integral of $f(p)$ over the range. (a) average lower, upper, and best-guess probabilities, (b) average distributions and (c) individual expert distributions for experts E1, E2 and E3 for the pre-session assessment.